

Manuscript category – Clinical Science**Title:**

Consensus generation of a minimum set of outcome measures for auditing glaucoma surgery outcomes - a Delphi exercise

Running title:

- A minimum glaucoma surgery outcomes dataset

Somner JEA¹

- FRCOphth, MD(Research) Vision and Eye Research Unit (VERU), Anglia Ruskin University, Cambridge Campus, Young Street, Cambridge, CB1 2LZ

Co-Authors

Ismail R² (PhD), Froud R³ (PhD), Azuara-Blanco A⁴ (FRCOphth), King AJ⁵ (FRCOphth)

Affiliations

- 1) Vision and Eye Research Unit, Anglia Ruskin University, UK
- 2) Royal Victoria Infirmary, Newcastle upon Tyne, UK
- 3) Warwick Medical School, University of Warwick, Coventry, UK and Kristiania University College, Oslo, Norway
- 4) Centre for Public Health, Queen's University Belfast, Belfast, UK
- 5) Nottingham University Hospital, Nottingham, UK

Corresponding Author

Anthony J King

Consultant Ophthalmologist
Nottingham University Hospital
Nottingham, NG7 2UH, UK

Email: anthony.king@nottingham.ac.uk

Word count manuscript: 1797

Word count abstract: 221 words

Purpose

To identify the key set of glaucoma surgery outcome measures considered most important and practical to collect by glaucoma specialists.

Methods

102 glaucoma specialists (57 members of the UK and Eire Glaucoma Society (UKEGS) and 45 members of the European Glaucoma society (EGS)) took part in an Online Delphi exercise. The RAND/UCLA appropriateness method was used analyse data from each round and generate a disagreement index.

Results

Participants agreed on 13 baseline data points and 12 outcomes that were considered important and practical to collect. For intraocular pressure (IOP) percentage reduction in IOP from baseline (last three IOP readings pre-op) and reduction below a specified target were considered important. For visual fields, change in a global visual field index e.g. MD and development of progression as assessed by linear regression were considered important. From a safety perspective, any visual loss resulting in a doubling of the minimal angle of resolution, loss of 5dB or more of visual field or development of advanced field loss (Hodapp Parrish Anderson Stage 4) were considered important. The importance of routinely using patient reported outcome measures (PROMs) was highlighted. Consensus suggested that outcomes of glaucoma treatments should be reported at one, five and ten years.

Conclusions

There was broad consensus on a minimum dataset for reporting the outcomes of glaucoma surgery and outcome measurement intervals.

Keywords

Glaucoma, Outcome Assessment (Health Care), Patient Outcome Assessment, Treatment Outcome

INTRODUCTION

Clinical audit is a mechanism to explore whether healthcare provision is in line with acceptable standards. This allows patients and care providers to know whether their service is doing well or whether improvements are required. Audit and its associated feedback mechanism is widely used as a strategy to improve professional practice. This has been an important aspect of care delivery in the UK for many years and is being more widely recognized as important in measurement of routine clinical outcomes (http://www.eugs.org/eng/outcomes_chair_action.asp) in glaucoma.

Healthcare professionals modify their practice when given feedback indicating that their clinical practice is inconsistent with a desired target¹ and consequently generating focus on providing high value care². In glaucoma audit it is essential to understand how well individuals and organisations deliver clinical care. This is important for both established treatment practices and for newly introduced novel treatments.

There is currently no established defined audit data set for evaluating clinical outcomes from glaucoma treatments and consequently no bench mark based on a defined and accepted core audit data set thus making comparisons between services and individuals impossible and preventing meaningful and effective audit from being achieved³.

Audit of cataract surgery outcomes⁴ has led to effective benchmarking standards being established and in the UK the NHS has established a national program of audit to benchmark outcomes for routine clinical practise (<https://www.hqip.org.uk/national-programmes/a-z-of-nca/>)

The aim of this study was through a consensus driven process to develop a core set of clinical outcomes which clinicians felt were important, reflected routine clinical care and were practical and easy to collect.

METHODS

A Delphi study, using a modified version of the Research AND Development (RAND)/University of California Los Angeles (UCLA) appropriateness method⁵ was administered via Survey Monkey. Potential important audit outcomes for glaucoma were identified from the literature^{6,7}. Panellists were asked to rate each potential outcome for importance for inclusion in a set of glaucoma surgery audit outcomes. For each potential outcome a 9-point scale was used to assess importance: 1 to 3 ('not important'); 4 to 6 ('indeterminate' or 'equivocal'); and 7 to 9 ('very important'). A free text field for each question was included to ensure we collected all opinions considered important by panellists.

A survey link was sent to 189 members of the UK and Eire Glaucoma Society (UKEGS) (Fifteen of these could not be delivered, leaving a total of 174 recipients) and attendees of the 2015 closed European Glaucoma Society (EGS) meeting in Tampere, Finland (130 members). The survey remained open for seven weeks, with weekly reminders sent to those who failed to respond.

Round One contained seven questions exploring consensus, including baseline characteristics, baseline IOP definition, treatment and safety outcomes, IOP and VF success measures and timing of outcomes evaluation. Round Two featured seven questions based on the refined outcomes of Round One and feedback from Round One panellists. The RAND/UCLA disagreement index was used to assess the degree of disagreement on rated items in both rounds⁵. An index of less than 1 is interpreted as no evidence of disagreement on an item. Participation in subsequent rounds was open to those who participated in a previous round and we judged consensus as present following a round with no disagreement on any of the rated items.

RESULTS

Sixty-four UKEGS members agreed to participate and responded. A further 38 EGS members agreed to participate and responded for a total of 102 participants. Seventy-five percent of panellists had previously carried out an audit of glaucoma outcomes with 21% having used an electronic medical record to facilitate the process. In Round Two, 84 consented to take part of whom 60 responded (71%, 42 UKEGS 18 EGS). The median time spent answering the Round One questionnaire was six minutes and 12 seconds and the Round Two questionnaire five minutes and 41 seconds.

The baseline characteristics the panellists agreed were important to collect are outlined in Table 1.

Disagreement index values were calculated separately for UKEGS and EGS responses – however there were no material differences in these values so the combined UKEGS/EGS DI responses are presented in the tables

The IOP value that should be used as the baseline was further explored in Round One (Table 2). There was broad agreement that selecting a baseline IOP is difficult, it would be useful to report both peak IOP off treatment and the mean of the last 3 IOP values before planned intervention on any topical medications (not diamox).

Table 3 lists what follow-up data was considered important and practical to collect, 16 outcome domains were ranked all of which were considered important except four.

Safety outcomes were explored in more detail. A loss of 3 lines of logMAR acuity was considered an important safety outcome by 84% of panellists with loss of perception of light suggested by 27% of panellists. Postoperative hypotony was also suggested in Round One and ranked in Round Two but there was disagreement between the groups as to whether this was an important safety outcome (median all 7, UK 8, EGS 5 – no disagreement all / UKEGS, EGS disagreement index 1.03).

Further evaluation of the importance of different visual field outcomes was undertaken in Round One and is shown in table 4

Panellists were asked which visual field parameters should be used to evaluate the success of treatments, a majority suggested that visual field outcomes should be measured in terms of both change in a global index and progression judged by linear regression. Interestingly there appeared to be a difference in emphasis here with UK

respondents favouring global measurement and EGS respondents favouring linear regression analysis. Two further visual field criteria were suggested for consideration by panellists in Round One and subsequently rated in Round Two: The loss of visual field by a fixed amount e.g. 5dB and the loss of visual field to a fixed level e.g. Hodapp Parrish Anderson criteria for advanced field loss (stage 4). Both were considered important visual field outcomes (median score for both items all 7, UKEGS 7, EGS 7 – no disagreement).

Success measures for IOP reduction were explored. The most popular overall measure was a percentage reduction from pre-treatment values followed by reduction to a clinician defined target IOP (Table 4).

Table 5 shows the follow-up time points important for outcome measurements to be collected. Twelve months and 5 years were the most popular choices. A further two time points were suggested for Round Two: 10 years (median all 7, UK 7, EGS 8 – no disagreement) and the last visit prior to death (median all 6, UK 6, EGS 7 – no disagreement).

DISCUSSION

Attempts are being made to use big data to improve the standard of care for patients with eye disease. Such projects work best if there are nationally or internationally agreed data sets⁸ and attempts have been made to achieve this for research outcomes.

Using a consensus approach to identifying the important measure for audit outcomes allows determination of what outcomes are most important or relevant to collect and ensures that important outcomes are not overlooked.

This study identified 13 baseline characteristics that were considered important to collect. Establishing a baseline data set is important as this will be used to define the risk/complexity of the cases undertaken which will be used for establishing weighted outcomes once sufficient data has been collected to allow this analysis to be undertaken.

Of the 12 treatment outcome measures identified as important many were perhaps unsurprising for a cohort of glaucoma specialists.

Regarding IOP, agreement on what should be recorded as the baseline IOP was achieved albeit requiring several different values to establish and this may be inconvenient in clinical practice. For IOP outcomes % reduction was the most popular and this is helpful as this is an objective assessment of IOP as opposed to target pressure measurement which may be different between clinicians because of subjective considerations⁹ and may also allow potential for gaming of outcomes. IOP is a commonly used outcome measure in glaucoma trials and is considered a surrogate for glaucoma progression, with lower IOP resulting in less visual field progression. However it only measures a single dimension of glaucoma care and it is acknowledged that IOP is an inadequate outcome measure used alone.¹⁰

Similarly for VF a consensus on measurement of VF outcomes is also helpful as multiple different measurement options exist for evaluating visual fields reported in clinical trials⁶ In this study panellists agreed on the easiest to collect data– global VF measures.

OCT is an important method of assessing structural glaucoma progression¹¹. However, it was not identified as an important outcome measure in this study. As part of the Delphi process respondents were asked to identify other measurements they thought should be included as outcome measures and despite this opportunity in each round, OCT was not considered one of these.

Several of the outcomes identified may be considered either bilateral (Visual impairment status, driving status, NEVFQ-25 vision related QoL, glaucoma related QoL) or generic health measurements such as generic QoL. While these are helpful perhaps in measuring a patient's glaucoma status as "a patient" they may not be helpful if reflecting the outcome of a particular intervention on "an eye" as the effects of this outcome will be diluted with influences from the fellow eye and general health condition. The inclusion of these measurements by a clinician cohort may reflect a growing awareness of the importance of considering all aspects of QoL in our treatment of patients. However these are not generally collected as routine data and therefore inclusion would add an extra burden to patients and resources if they were to be collected.

The results suggest some consensus among experts that outcomes should be measured at 12 months, which would be generally considered at the end of the active phase of many surgery interventions and at 5 years which, may provide more meaningful data on the life-time outcomes of an intervention.

One limitation of this study is that patients as the service users were not involved in the process of selecting the outcomes, which should be included in an audit of glaucoma care provision. This requires further investigation and these findings should be combined with the outcomes of clinician preferences to form a composite outcomes dataset.

Although consensus was reached on many elements within this study, the existence of consensus from a Delphi process does not mean that the 'correct' answer has been found. The participants were opportunistically chosen through membership of the EGS and UKEGS – however, we consider that a meaningful range of expert views were captured across our 60 completing panel experts.

Identification of important outcomes in cataract surgery has led to a valid, evidence based benchmarking process⁴ which is now used to define acceptable levels of posterior capsule rupture rates in the UK.

We anticipate that the information developed in this research will prove helpful to those seeking to undertake audits of glaucoma surgery. Through identifying a minimal dataset the audit process will be simplified and maximum benefit achieved. It is hoped that the results of collection of large amounts of audit data on glaucoma outcomes will not only allow benchmark setting but also risk stratification for outcomes measured.

Acknowledgements

Thanks to UKEGS and EGS for providing access to their mailing lists, to all the consultants who took the time to fill out the questionnaire and to Professor Anja Tuulonen for critical reading of the manuscript and helpful suggestions. RF is a director and shareholder of Clinvivo Ltd, a University of Warwick spin-out company that provides services including design and administration of Delphi studies. These services were not used in this study and notwithstanding the disclosure RF declares no conflicts of interest. None of the other authors has any competing financial interests in relation to the work described.

Compliance with Ethical Standards:

Funding: No funding was received to support this study.

Conflict of Interest:

Somner declares that he has no conflict of interest

Ismail declares that she has no conflict of interest

Azuara-Blanco declares that he has no conflict of interest

King has received honoraria from Allergan UK for lectures performed

Froud declares that he has no conflict of interest. **H is a director and shareholder of Clinvivo Ltd, a University of Warwick spin-out company that provides services including design and administration of Delphi studies. These services were not used in this study**

Ethical approval: This article does not contain any studies with human participants performed by any of the authors.

Informed consent: Informed consent was not required for this survey

References

1. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. (2012) Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane database Syst. Rev.* ;(6): CD000259.
2. Porter ME. (2010) What is value in health care? *N. Engl. J. Med.* ;363(26): 2477–81.
3. Rotchford AP, King AJ. (2010) Moving the goal posts definitions of success after glaucoma surgery and their effect on reported outcome. *Ophthalmology.* ;117(1): 18–23.
4. Mahmud I, Kelley T, Stowell C, Haripriya A, Boman A, Kossler I, et al.(2015) A Proposed Minimum Standard Set of Outcome Measures for Cataract Surgery. *JAMA Ophthalmol.* ;133(11): 1247.
5. Fitch K. (2001) The RAND/UCLA appropriateness method user's manual. Rand Corp.Santa Monica
6. Ismail R, Azuara-Blanco A, Ramsay CR. (2014) Variation of clinical outcomes used in glaucoma randomised controlled trials: a systematic review. *Br. J. Ophthalmol.* ;98(4): 464–468.
7. Ismail R, Azuara-Blanco A, Ramsay CR. (2016) Consensus on Outcome Measures for Glaucoma Effectiveness Trials. *J. Glaucoma.* ;25(6): 539–546.
8. Jackson TL, Donachie PHJ, Johnston RL et al. (2016) Interpreting Database Studies. *JAMA Ophthalmol.* ;134(1): 86–87.
9. Solano-Moncada F, Dymerska M, Jefferys JL, Quigley HA. (2016) Adherence With the Use of Target Intraocular Pressure for Glaucoma Patients in a Large University Practice. *J. Glaucoma.* ;25(4): e424-32.
10. Spaeth GL. (1995) Proper outcome measurements regarding glaucoma: the inadequacy of using intraocular pressure alone. *Eur. J. Ophthalmol.* ;6(2): 101–5.
11. Tatham AJ, Medeiros FA.(2017) Detecting Structural Progression in Glaucoma with Optical Coherence Tomography. *Ophthalmology*;124(12S):S57-S65.